



**University of  
Zurich**<sup>UZH</sup>

**Zurich Open Repository and  
Archive**

University of Zurich  
University Library  
Strickhofstrasse 39  
CH-8057 Zurich  
[www.zora.uzh.ch](http://www.zora.uzh.ch)

---

Year: 2018

---

## **Apologies in the History of English: Evidence from the Corpus of Historical American English (COHA)**

Jucker, Andreas H

**Abstract:** This paper explores two different methods of tracing a specific speech act in a historical corpus. As an example, the development of apologies is investigated in the two hundred years covered by the Corpus of Historical American English (COHA, 1810-2009). One method retrieves apologies through their typical illocutionary force indicating devices (IFIDs), such as sorry, excuse, apologise and pardon, while the other retrieves passages in which apologies are explicitly mentioned (metapragmatic expression analysis). Both methods require a considerable amount of manual analysis of retrieved hits, which has to be verified through elaborate inter-rater reliability testing. The searches are restricted to fictional texts because they show a greater frequency of apologies than the alternative genres available in COHA, and they often allow the identification of behaviour as apologetic because it is discursively described as such by the fictional characters or the narratorial voice. The results show that the frequency of apologies increased considerably throughout the period covered by COHA. In the earliest period the IFID sorry was no more frequent than pardon and forgive. In the most recent period its frequency has multiplied almost six-fold and is more than three times larger than all the others taken together. The metapragmatic expression analysis allows an analysis of the development of strategies used to perform apologies. IFIDs have become more important while Taking on Responsibility and Explanation receded somewhat in their frequencies. On the basis of these results it is speculated that the force of apologies has decreased. What used to be sincere requests for exoneration has in many cases turned to token displays of regret.

DOI: <https://doi.org/10.1007/s41701-018-0038-y>

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-151590>

Journal Article

Accepted Version

Originally published at:

Jucker, Andreas H (2018). Apologies in the History of English: Evidence from the Corpus of Historical American English (COHA). *Corpus Pragmatics*, 2(4):375-398.

DOI: <https://doi.org/10.1007/s41701-018-0038-y>

## **Apologies in the history of English: Evidence from the *Corpus of Historical American English* (COHA)**

Andreas H. Jucker, University of Zurich

### **Abstract**

This paper explores two different methods of tracing a specific speech act in a historical corpus. As an example, the development of apologies is investigated in the two hundred years covered by the *Corpus of Historical American English* (COHA, 1810-2009). One method retrieves apologies through their typical illocutionary force indicating devices (IFIDs), such as *sorry*, *excuse*, *apologise* and *pardon*, while the other retrieves passages in which apologies are explicitly mentioned (metapragmatic expression analysis). Both methods require a considerable amount of manual analysis of retrieved hits, which has to be verified through elaborate inter-rater reliability testing. The searches are restricted to fictional texts because they show a greater frequency of apologies than the alternative genres available in COHA, and they often allow the identification of behaviour as apologetic because it is discursively described as such by the fictional characters or the narratorial voice. The results show that the frequency of apologies increased considerably throughout the period covered by COHA. In the earliest period the IFID *sorry* was no more frequent than *pardon* and *forgive*. In the most recent period its frequency has multiplied almost six-fold and is more than three times larger than all the others taken together. The metapragmatic expression analysis allows an analysis of the development of strategies used to perform apologies. IFIDs have become more important while Taking on Responsibility and Explanation receded somewhat in their frequencies. On the basis of these results it is speculated that the force of apologies has decreased. What used to be sincere requests for exoneration has in many cases turned to token displays of regret.

Keywords: Apologies, speech acts, diachronic corpus analysis, COHA, history of American English

### **1 Introduction**

The following extracts all contain apologies. They are drawn from very different sources and very different time periods.

- (1) “I pray your pardon for disturbing your repose.” (COHA 1831)
- (2) “I cry you mercy,” said I, “for mistaking your age; but it matters little.” (COHA 1849)
- (3) I beg your forgiveness a thousand times for not having sooner sent my apologies. (COHA 1866)
- (4) We pay them for their deeds, literally. (soz for the off topic rant) (GloWbE US 2001)

- (5) No, I will not give you a synonym. Go home and look up the word! Oops, sorry.  
My bad. Go on your smart phone and Google the word. (COCA 2013)

It is clear that these apologies differ according to several dimensions. The speakers or writers use a variety of very different linguistic means to apologise, the offences for which they apologise appear to differ in their severity, and the seriousness of the apologisee also seems to differ considerably. Some of the apologies come across as heartfelt expressions of regret about what has happened while others appear to be no more than token or indeed ironic expressions of regret. The examples are spread out over almost two hundred years, and they are in some ways typical of their periods as I will show below. Nevertheless, there has always been a fair amount of variation in the linguistic resources available for apologies.

In the extant literature it appears that it is generally unproblematic and straightforward to identify apologies (but see Lakoff 2001, who draws attention to the fuzzy nature of apologies). We certainly recognise an apology when we see one, at least most of the time. For some researchers the method of investigation itself guarantees that what they are dealing with are actually apologies. Experimental methods, such as discourse completion tasks or role plays, are designed to elicit apologies, and hence the outcome appears to be unproblematic instantiations of apologies.

The literature on apologies seems to agree on the basic outlines of the functional profile of apologies. There are four crucial elements that are generally mentioned in this respect. An apology needs an offence that has been committed, an offender who takes responsibility for the offence, someone who feels offended or is believed by the offender to possibly feel offended, and a display of recognition of and regret for the offence (see, for instance, Holmes 1990: 159; Olshtain 1989: 156-157; Deutschmann 2003: 46; Lutzky and Kehoe 2017a: 28 and countless others). However, as is also regularly noted, all four aspects are not as clear-cut as might appear at first sight.

We may notice that this is a relatively wide interpretation of what an apology is. Robinson (2004), for instance, restricts the term to what he calls “explicit apologies” in order to distinguish them from “other offense-remedial-related actions, such as accepting blame (e.g., *It’s my fault*), promising forbearance (e.g., *I promise it won’t happen again*), requesting forgiveness (e.g., *Forgive me*, and *I beg your pardon*), and requesting to be excused (e.g., *Excuse me*) and pardoned (e.g., *Pardon me*)” (Robinson 2004: 292; see also Goffman 1971 and Owen 1983). Lakoff (2001: 201, 205) also draws attention to the difficulty to recognise apologies because they range from canonically explicit formulations to ambiguously indirect ones, and they can be hard to distinguish, for example, from explanations, excuses and justifications.

Corpus searches have generally focused on a small range of apology expressions, such as *sorry*, *apologize*, *excuse*, *forgive* and so on, which seem to unequivocally signal the presence of an apology. Deutschmann (2003: 36) famously claimed that “apologising tends to be accompanied by a limited set of easily identifiable routine formulae. Of course it is theoretically possible to apologise without saying *I’m sorry* or *excuse me* but research has shown that this is rarely the case in English”. As evidence he cites Meier (1998)<sup>1</sup>, who in

---

<sup>1</sup> In fact, he cites Meier (1994) but such a title does not exist in his references. The intended reference must be Meier (1998).

turn quotes, for instance, Holmes (1990) and Olshtain (1989). Holmes (1990: 167) reports 96 per cent of apologies in New Zealand English to contain an IFID, and Olshtain (1989: 164) 75 per cent in Australian English. However, Holmes used the diary method and Olshtain discourse completion tasks (DCTs). It is an entirely open question whether any of these rates bear any resemblance to apologies attested in a variety of genres in corpora.

Historical investigations inescapably have to rely on corpus material. Discourse completion tasks or the diary method cannot be used. However, in the following I want to show how a combination of different search methods can lead to a more comprehensive understanding of the historical development of a specific speech act. Thus, I will combine a search for apology IFIDs in the *Corpus of Historical American English* with a search for metapragmatic expressions, which are expressions that explicitly mention a pragmatic act, in this case the speech act of apology. The term used in this investigation is *apolog\**, which represents the various derivations of *apologize*, such as *apology*, *apologetic*, *apologetically*, etc.). Such a search retrieves passages in which apologies are explicitly talked about, either by the interactants or by a narratorial voice. The search results require a detailed inspection and analysis of every single hit, or a representative sample of hits, but they offer a much more detailed perspective on the development of apologies than a corpus-analysis based on IFIDs alone would be able to.

## 2 Relevant literature

Apologies have received a great deal of attention from various different theoretical perspectives, and investigations have used a variety of methods. Research into the history of apologies is less frequent but some work already exists. Kohnen (2017), for instance, traces the speech acts of boasting and apologising in Old English and argues that their status then was very different from what it is today. Today, boasting, and in particular too much boasting, may often seem inappropriate, but in the Anglo-Saxon world of warrior heroes it seemed entirely appropriate to show off your own achievements, enhance your own face and boast about your successes. Apologies, on the other hand, appear to be entirely appropriate and indeed expected in certain present-day situations, but for the Anglo-Saxon hero there was no need to apologise. An offence was a transgression that required retribution and punishment but no apology (Kohnen 2017: 305, 313). Indeed, there was no speech act verb “to apologise” in Old English. When Kohnen searched for relevant lexemes expressing sad feelings, regret, excuse and forgiveness (on the analogy of Present-day English apologies including expressions, such as *sorry*, *regret*, *excuse* and *forgive*), he found acts of penitence and repentance but not of apology. Extracts (6) and (7) are typical examples.

- (6) ofhreoweþ me swa hwæt swa ic dyde oþþe gepohte ongen bebodu þine & ongen þinne haligan willan. (ArPrGl 1 (Holt-Campb) C23.1)  
‘I repent whatever I did or thought contrary to your commandments and contrary to your holy will.’
- (7) Miltsa me, drihten, hæl mine sawle, forðon me hreoweð nu þæt ic firene on ðe fremede geneahhige. (PsFr A24)

‘Take pity on me, Lord, cure my soul, because I now repent that I committed wicked deeds against you in abundance.’ (Kohnen 2017: 314)

In extract (6), the speaker expresses repentance to God for his sins or what is very generally described as “whatever I did or thought contrary to your commandments and contrary to your holy will”. This is very different from an apology for a particular offence. In a similar fashion the speaker of extract (7) expresses remorse for an equally vague set of sins or “wicked deeds”. On the basis of such examples, Kohnen (2017: 316) argues that acts of penitence in a Christian context constituted a kind of pre-apology in the larger Anglo-Saxon world in which there was no room for apologies as we know them today.

Jucker and Taavitsainen (2008) investigated the history of apologies in the Renaissance period (1500-1660) in a corpus containing fiction and drama texts from this period. They found that apologies were less routinized than they are today. A simple apology expression, such as *sorry* or *pardon* did not seem sufficient to apologise. People regularly added terms of address, explanations and phrases, such as “I beseech you” or “I pray you”. Extracts (8) and (9) are relevant examples.

- (8) No more my Lord at this time, I am sorry that I have given you such cause of grieffe, thus by recounting so lamentable a state, to renew your passed griefes. But comfort good King, when Tydes be at the lowest, they spring againe. (LION; Anon., Marianvs (c. 1641), page 139)
- (9) Honest man, I pray you pardon me, if I say any thing that may offend you; I am sorie to see the euil that is towards you: you haue bene very mery, but I feare, you will neuer be so againe in this company: for I see in your eyes a spirit of madnesse, which will very speedily bring you to your vnhappy ende: for indeede, within this houre, you will hang your selfe in the stable, vpon one of the great beames: (LION; Anon., Pasquils Iestes (1609), page 41) (Jucker and Taavitsainen 2008: 238, 240)

They notice that in their data apologies for speaking offences, for lack of decorum, for being too direct, rude or impolite are particularly frequent. Without being able to rely on statistics they also speculate that addressee-oriented apologies, which ask the addressee for forgiveness (*pardon*, *excuse*, *forgive*) appear to be more frequent than speaker-oriented apologies, which express the speaker’s feeling of remorse (*sorry*). They tentatively conclude that Renaissance apologies are more like requests for the addressee’s generosity to forgive or overlook the offence while present-day apologies appear to express the speaker’s remorse and refrain from imposing on the addressee to forgive the offence.

For Present-day English the research situation is much better. Apologies have been investigated from many different perspectives and with a variety of methods, ranging from philosophical (e.g. Searle 1969) and conversation analytical approaches (e.g. Robinson 2004; Heritage and Raymond 2016) to experimental (Blum-Kulka et al. 1989a; Trosborg 1987, 1995) and corpus-based approaches (e.g. Aijmer 1996; Deutschmann 2003; Lutzky and Kehoe 2017a, b).

Blum-Kulka et al. (1989a), for instance, included apologies in their cross-cultural speech act investigation. Requests and apologies were considered to be particularly suitable speech acts for this investigation because they constitute specific face threats in the sense of Brown and Levinson (1987). Requests are seen as threats to the addressee’s negative

face because they impose on the addressee and ask him or her to do something that they perhaps would not have done otherwise. Apologies, on the other hand, are seen as a threat to the speaker's own positive face since the apologizer admits to having done something untoward and thus damages his or her own positive face. The investigation was carried out with the help of discourse completion tasks (DCTs) that were given to large numbers of students in different cultural and linguistic settings. They were asked to imagine a range of everyday situations and write down what they would say in such a context. Extracts (10) and (11) illustrate two such situations, which were designed to elicit an apology.

(10) *At the College teacher's office*

A student has borrowed a book from her teacher, which she promised to return today. When meeting her teacher, however, she realizes that she forgot to bring it along.

Teacher: Miriam, I hope you brought the book I lent you.

Miriam:.....

Teacher: OK, but please remember it next week

(11) *In the lobby of the university library*

Jim and Charlie have agreed to meet at six o'clock to work on a joint project.

Charlie arrives on time and Jim is half an hour late.

Charlie: I almost gave up on you!

Jim:.....

Charlie: O.K. Let's start working.

(Blum-Kulka et al. 1989b: 14, 1989a: 274)

Since the publication of this research more than thirty years ago, discourse completion tasks have come in for extensive criticism. They have been condemned both on theoretical and on practical grounds. The responses that the participants in experiments write down are not what they would actually say in such situations because the task is artificial and it asks participants to put into writing what they would normally perform orally. In a normal situation, a speaker does not know in advance how the addressee is going to react but in a DCT the reaction is already given. The task is printed in such a way that it requires exactly one turn from the apologizer where in reality a sequence of exchanges might be needed, and the space is limited to relatively short apologies. For an up-to-date assessment of discourse completion tasks and an overview of the extant criticisms and defences of the method see Ogierman (2018).

In order to compare the realization of the speech acts under investigation Blum-Kulka and her team developed a very detailed categorization scheme consisting in the case of apologies of five main strategies and a large number of more detailed sub-strategies. The main strategies were Illocutionary Force Indicating Device (IFID), taking on responsibility, explanation or account, offer of repair and promise of forbearance. These strategies can be used singly or in combination. Example (12) illustrates a case in which all main strategies are being used.

(12) I'm sorry (IFID), I missed the bus (RESPONSIBILITY), and there was a terrible traffic jam (EXPLANATION). Let's make another appointment (REPAIR). I'll make sure that I'm here on time (FOREBEARANCE)

(Blum-Kulka et al. 1989a: 290; see also Olshtain 1989)

Trosborg (1987, 1995) also used an experimental situation to elicit apologies. She video-recorded pairs of interactants in role plays. She asked the participants to act out a particular situation which was designed to provoke an apology. She also carefully monitored the variables social distance and dominance in these situations in order to find out how they influence the way in which apologies are realized, and she compared three different groups of speakers; native speakers of Danish, native speakers of English, and native speakers of Danish speaking English.

In more recent research, people have turned away from elicitation techniques and relied mostly on corpus-based methods or on collections of apologies drawn from large sets of transcribed spoken language. Deutschmann (2003), for instance, set out to investigate apologies in a large corpus of spoken texts. He chose the 5-million-word spoken part of the *British National Corpus* and identified 3,070 explicit apologies within half a million utterances, which yields a frequency of about 60 apologies per 100,000 words or per 10,000 utterances. He searched for the apologies on the basis of a list of seven illocutionary force indicating devices, i.e. *sorry*, *pardon*, *excuse*, *afraid*, *apologise*, *forgive* and *regret*, as well as a number of modifications of these seven. He claims that apologies without any of these are rare (Deutschmann 2003: 36; see section 4 for a discussion of this claim). It turns out that *sorry* is by far the most frequent of these IFIDs, accounting for 1,820 apologies. *Pardon* accounts for about half as many (815) and *excuse* for roughly a sixth (320). The other four together account for the remaining 115 apologies (Deutschmann 2003: 51). Deutschmann then proceeds to analyse this set of 3,070 apologies in great detail, looking both at the syntactic complexities of the individual formulations and at the range of offences for which they are used. He proposes a detailed catalogue of offence types and provides precise statistics as to the frequency of each type in his data. He finds that hearing offences, such as not hearing, not understanding or not believing one's ears, constitute the largest group accounting for more than 30 per cent of all cases in his data (Deutschmann 2003: 64). Lack of consideration accounts for about half as many cases. This is followed by breach of consensus, talk offences, misunderstandings and mistakes, and breach of expectations each account for about ten per cent or slightly less. The remaining categories, social gaffes, requests and accidents are somewhat less frequent still, which leaves a surprisingly small rest of unidentified cases (4 per cent).

On the basis of the difficulties of categorising the data in the *Corpus of Historical American English* encountered in this research project, the great precision of a very fine-grained analysis appears remarkable and admirable. As I will outline below, in our data it proved to be difficult and often impossible to reach a sufficient level of inter-rater agreement even with much simpler categorisation schemes.

Drew et al. (2016) introduce a collection of papers based on the same set of 200 apologies systematically collected from all available telephone call corpora by Gail Jefferson in 2003. According to the editors the only research method that counts as empirical consists of “research based on analysis of naturally occurring instances of apologies in interpersonal interactions” (2016: 3). Heritage and Raymond (2016), to pick out one example from the collection, investigate Goffman's (1971: 116) claim that there is a proportional relation between the offence and the formulation of the apology. They distinguish between what they call “local” and “distal” offences, the former of which generally have to do with speaking and hearing problems in the context of the apology itself while the latter relate to past or future conduct. Moreover, this distinction is related

to whether or not an offence needs to be made explicit. In the case of local problems, the problem was immediately available to both participants and generally did not need to be mentioned, while distal problems distinguished between problems that were and that were not available to both participants (Heritage and Raymond 2016: 6). They found that distal offences generally required more complex forms of apologies.

Robinson (2004: 295) likewise uses a collection of apologies drawn from a corpus of carefully transcribed naturally-occurring interactions including telephone calls, dinner-table conversations and other social interactions. He focuses on the sequential positioning and in particular the adjacency pair organization of apologies.

More recently Lutzky and Kehoe (2017a, b) have tackled the problem of how to track apology IFIDs in large corpora on the assumption that a micro analysis is not possible or practicable if a corpus is too large, as for instance in the *Birmingham Blog Corpus* (BBC), which they use for their investigation. This corpus contains 600 million words from 2000 to 2010. They solve this problem by establishing collocational profiles for each apology expression and by comparing shared and unique collocates across these profiles. This helps them to identify non-apology uses of apology expressions and to uncover additional apology expressions, something that was not possible with Deutschmann's (2003) method. In particular they identify *oops* and *whoops* and a number of spelling variants of these plus *my bad* as additional apology expressions or IFIDs. Thus, they use this method as a way of bridging the gap between analysing individual richly contextualised examples and analysing very large numbers of examples.

### 3 Corpus analysis

An investigation of the historical development of a speech act must necessarily rely on corpora, which are searched either electronically or manually. Experimental methods cannot be used because native speakers of earlier periods are not available. Historical corpora are now available for various historical periods and languages but all of them are severely restricted in the types of material that they contain. The compilers of historical corpora have to focus on those materials that are available at all, that are available in sufficient quantity and that exist in comparable form over a longer period of time. One easily available historical corpus is the *Corpus of Historical American English* (COHA), which currently contains some 400 million words and covers the genres fiction, magazines, newspapers and non-fiction books. Conversations as such or other forms of spoken language are not included except for their representation in the four genres mentioned above. Thus, it is not possible, on the basis of this corpus, to investigate the history of everyday spoken interaction.

Preliminary investigations soon revealed that a large majority of all the apologies in the corpus are attested in the fictional material. They occur in the other genres as well but at a much lower frequency. For this reason, I decided to focus entirely on the fiction part of the corpus, which makes the investigation more focused. Fictional representations are not offered here as a substitute for everyday spoken interaction. They merely show how authors chose to represent spoken interactions in their works of fiction, and all the claims about forms, functions and developments of apologies proposed in this paper, are in fact claims about how apologies are represented in fictional texts, and it is, of course, possible that



some of the developments of apologies reported in this paper reflect changes in literary styles.

In fact, fictional texts offer some analytical advantages over transcriptions of spoken interaction. Fictional texts generally describe self-contained worlds. The depicted characters do not have a life outside of this fictional world, and, therefore, the motivations for all their behaviour must be sought within the limited universe of the fictional world, while interactants in everyday spoken interaction bring with them a large amount of life experience that remains hidden to the analyst and cannot be brought to bear on a precise understanding of how and why a specific speech act is being used. In addition, fictional texts often have a narratorial voice, which situates and explains the behaviour of the fictional character, as, for instance, when a character is said to nod apologetically. In an everyday conversation, a nod may be perceived by the bystanders and the analyst to be apologetic, but in all probability the behaviour may well be ambiguous and indeterminate. The narratorial voice makes its status more explicit and thus provides the researcher with an extra analytical handle that ties in very directly with today's understanding of speech act values to be partly a matter of the linguistic resources used to perform them and partly a matter of discursive negotiation between the interlocutors. In this case the author of a work of fiction can use the narratorial voice or the interactions by the characters to discursively specify intended speech act values.

The *Corpus of Historical American English* covers a two-hundred-year period from 1810 to 2009. It is split up into individual decades, which allows the researcher to trace items of interest in their decade-by-decade development. In many cases, however, this turns out to be too fine-grained, especially if certain constructions are relatively rare. The early decades of the corpus contain less data in comparison to more recent decades, especially if only the fiction section is considered. In fact, all the decades before 1900 contain less than 10 million words in the fiction sections, and the first decade (1810-1819) only a little more than half a million. For this reason, I decided to adopt a somewhat wider perspective and split the entire period covered by COHA into four subperiods of 50 years each. Table 1 gives an overview of the subcorpora used for this study.

A 1810-1859	24,147,783
B 1860-1909	45,207,474
C 1910-1959	52,147,145
D 1960-2009	55,487,202
Total	176,989,604

Table 1: Number of words per subperiod of COHA (fiction only)<sup>2</sup>

The first period contains just over 24 million words of fiction. The other three periods are of roughly equal size and about twice as large as the first one.

<sup>2</sup> The word counts reported here were established via the Dependency Bank 2.0 of the English Seminar of the University of Zurich. The word counts deviate to some extent from that obtained through the official COHA website (see Data Sources below).

### 3.1 Corpus-based speech act analysis

Speech acts cannot be searched for directly. They are entities that have functional definitions which cannot be transformed in any straightforward way into search strings. In earlier work, Irma Taavitsainen and I have proposed three different solutions to this problem (Jucker and Taavitsainen 2013: ch. 6, 2014: 258-261). All of them have their shortcomings, but they do help to retrieve at least a subset of relevant items from corpora.

The first solution uses a list of known illocutionary force indicating devices (IFIDs) to retrieve relevant speech acts. Thus, *please* will retrieve requests and *sorry* will retrieve apologies. But they will only retrieve some of the desired speech acts (limited recall) and some of the retrieved hits will turn out to be something else (limited precision). This is the method used by Aijmer (1996), Deutschmann (2003) and Jucker and Taavitsainen (2008) for apologies. It has also been used, for example, for requests (Aijmer 1996) or for greetings and farewells (Jucker 2017). In addition to the problems of precision and recall, this solution does not work for speech acts which are not sufficiently conventionalised and do not have reasonably frequent IFIDs.

The second solution, therefore, searches for specific patterns that are typical for a certain speech act without directly functioning as an IFID. Compliments, for instance, have been claimed to be regularly realised in a very limited set of syntactic patterns (Manes and Wolfson 1981). The most frequent pattern, for instance, is said to consist of a noun phrase, a linking verb, such as *is* or *looks*, an optional intensifier (*really*, *very*) and a positive adjective (e.g. “Your skin looks amazing.” COCA, NBC Today Show, 2014). The second pattern consists of the first person singular pronoun *I*, an optional intensifier, a linking verb such as *love* or *like* and a noun phrase (e.g. “I really love your piano playing.” COCA, NPR: Fresh Air, 2017). According to Manes and Wolfson, the three most frequent patterns already account for 85 per cent of their entire data set of 686 compliments (Manes and Wolfson 1981: 120). These patterns together with the less frequent ones proposed by them were taken as the starting point for the corpus investigation by Jucker et al. (2008). They transformed the syntactic patterns into search strings and used these to retrieve compliments from the *British National Corpus*, and they found that the frequencies of the different patterns differed significantly from those reported by Manes and Wolfson.

The third solution takes a different approach and searches not for a particular speech act itself but for passages in which this speech act is the object of conversation, that is to say the searches retrieve metapragmatic expressions, i.e. expressions that are used to talk about a pragmatic entity such as a speech act. Jucker and Taavitsainen (2014) used this method to trace the history of compliments in American English. The retrieval of passages containing a potential metapragmatic expression is obviously only the first step in the analysis. The actual analysis, then, consists of a manual inspection of the retrieved hits or of representative subsets of retrieved hits.

As pointed out above, apologies regularly make use of an illocutionary force indicating device, such as *sorry*, *pardon* or *excuse*, but beyond that they do not appear to use conventionalized patterns that might be used as search strings. In this case, therefore, only the first and the third method can be used, and it is the aim of the following analysis to show how a combination of these methods can provide a better understanding of the developments than a reliance on IFIDs alone. Thus, I will first present an analysis of searches that were based on apology IFIDs along the lines of previous research (Aijmer

1996, Deutschmann 2003) but applied to historical data and – in contrast to Jucker and Taavitsainen (2008) – with a focus on a different time period and with a view to get not only qualitative and impressionistic but also quantitative results accounting for actual diachronic developments of apologies. In a subsequent step I will then adopt the third method and explore the potential of a metapragmatic expression analysis with the same historical data of American English from 1810 to 2009.

### 3.2 Corpus analysis based on IFIDs

In a first step the distribution of the most common illocutionary force indicating devices was established across the four half centuries of the *Corpus of Historical American English* (COHA). However, the raw frequencies of these expressions do not give an accurate indication of the frequency of apologies because many of these expressions have homonyms with uses outside of apologies. What I was interested in in this step was the frequency of these expressions when they are actually being used in the context of a performative apology. Extracts (13) to (17) give an impression of the range of uses for the expression *sorry*.

- (13) and the man turned to her in surprise. “I’m sorry, I shouldn’t have said that,”  
Eileen said. (COHA, fic, WheelLove, 1970)
- (14) so after a minute I said, “Mother, I’m sorry. I didn’t mean that. Please don’t cry.  
Please.” (COHA, fic, Harpers, 1963)
- (15) “And a sorry thing that will be, from this day forward!” (COHA, fic,  
SpeaksNightbird, 2004)
- (16) He maybe was a shit but I felt sorry for the guy. (COHA, fic, FantasySciFi, 1994)
- (17) “I thought,” said Grace, “the sun must look very jolly in his red silk night-cap,  
only I was sorry you forgot to tell what he had for breakfast.” (COHA, fic,  
DottyDimpleAtHer, 1867)

In extracts (13) and (14) the fictional speakers issue an apology and express regret for their own actions. In (15) *sorry* is used as an adjective meaning something like ‘pitiful’ or ‘unpleasant’. In (16) the speaker expresses pity for somebody else, and in (17) the speaker expresses remorse for somebody else’s actions.

For this exercise only performative instances of apology expressions, as illustrated by extracts (13) and (14) for *sorry*, were considered. In some cases, apology expressions, e.g. *apologise* or *excuse*, turned out to retrieve a very large number of hits with very poor precision. In these cases, the search string was modified to increase the precision with a necessary slight reduction of recall. Table 2 lists the regular expressions that were used to retrieve the individual apology expressions.

Lexical head	Regex
sorry	sorry
excuse	excuse (me my myself us our ourselves)

apologise	(I let me we let us) apologi[sz]e
forgive	forgive (me my us our)
pardon	pardon (me my us our) your? pardon
regret	(I we) regret
afraid	(I am we are) afraid

Table 2: Regular expressions used to retrieve apology expressions (IFIDs) (expressions in brackets separated by vertical bars indicate lists of choices)

The number of hits retrieved with the regular expressions in Table 1 was still too large for manual inspection in all cases. It was, therefore, necessary to draw representative samples to assess the rate of performativity for each expression. A preliminary inspection of the data clearly suggested that this rate of performativity was not constant throughout the period covered by COHA. It was, therefore, necessary to assess the performativity individually for each expression for each of the four half-centuries. For this, two trained coders first independently coded identical samples of 100 hits for each expression in order to ascertain the level of inter-rater agreement. In most cases the agreement was well over 80 per cent, except for *regret*, where only a level of 76 per cent was reached. On this basis, the two coders then coded one hundred hits for each expression and each half-century to determine the performativity of this expression for this particular period. In the case of *sorry*, for instance, there was a very marked increase of the rate of performativity over the four periods, from 36 and 39 in the first two periods to 48 and 67 in the third and fourth period. This rate of performativity was then used together with the number of originally retrieved hits per expression and period to calculate an approximate number of instances in which the expressions were actually used in an apology over the entire fiction material for each period. Figure 1 shows the result of this investigation.

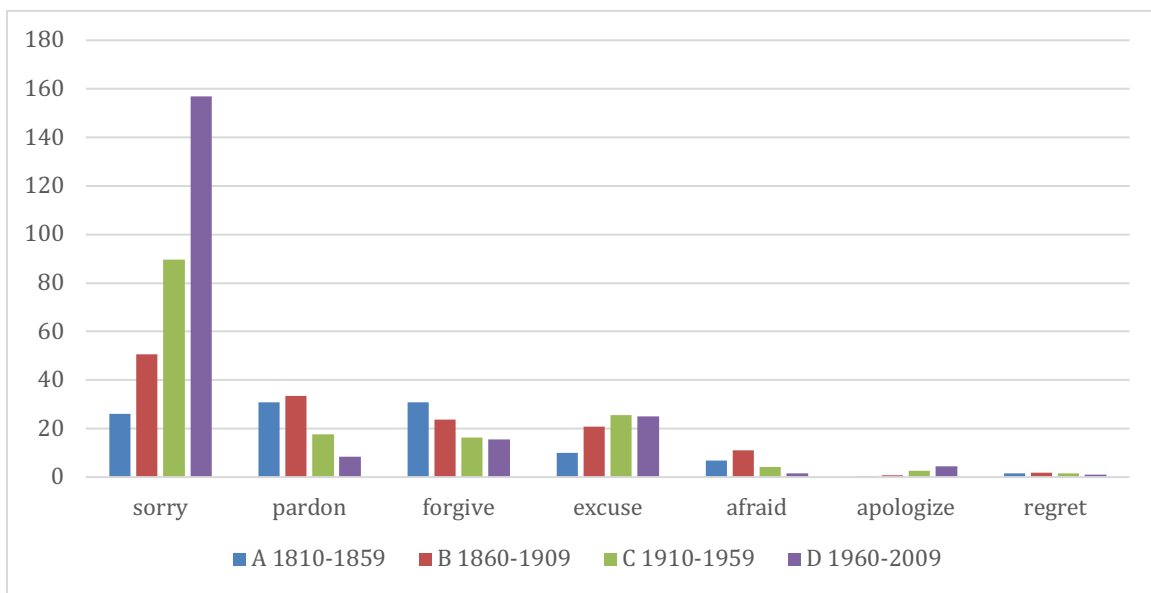


Figure 1: Frequency of performative apology expressions (IFIDs) in the four half-centuries in the fiction section of COHA (per million words)

The individual figures represented in Figure 1 are, of course, no more than reasonable approximations. The recall was not perfect because it was restricted to make it more manageable; the inter-rater reliability tests, as pointed out above, have shown that agreement on the performativity is high but not perfect; and it is, of course, possible that the rate of performativity of individual expressions varies also throughout an individual period not just across periods. However, the trends appear to be clear. In the first half-century, the three apology expressions *sorry*, *pardon* and *forgive* each had a roughly equal frequency of 25 to 30 instances per million words. Other expressions were negligible in comparison. Across the two centuries both *pardon* and *forgive* diminish to less than ten and about 15 instances, respectively, while *sorry* clearly takes over and shows a steady increase over the half-centuries from 26 to 50 and 90, and in the most recent period to over 150 instances per million words. *Excuse* and *apologise* also increase over time but at a very modest level. *Excuse* rises from ten to 25 instances per million words and *apologise* from 0.1 to about 5 instances.

In the first half-century, therefore, the three expressions *sorry*, *pardon* and *forgive* were used almost equally to perform apologies. Over the two centuries this gradually changed to a situation in which *sorry* accounts for about three quarters of all apology expressions, and the second most frequent one, *excuse*, to only just above ten per cent. Figure 2 shows this development across the four half-centuries.

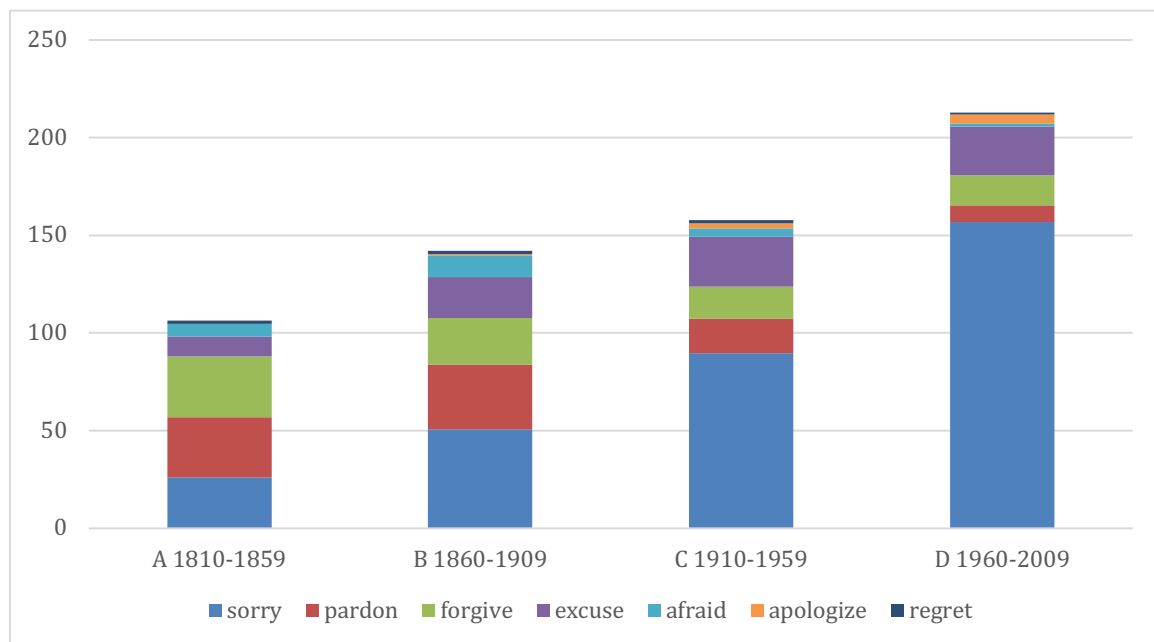


Figure 2: Combined frequency of performative apology expressions (IFIDs) in the four half-centuries in the fiction section of COHA (per million words)

Again, it must be borne in mind that these statistics do not cover all apologies. They only cover those apologies that contain an apology expression (an IFID), and they only cover apologies that are recorded in fictional texts. In the next step, therefore, I shift the analysis to passages in which narrators or fictional characters explicitly talk about apologies.

### 3.3 Corpus analysis based on metapragmatic expressions

A second way of locating specific speech acts in a corpus, as pointed out in section 5.1 above, consists in the retrieval of metapragmatic expressions. This method has also been called analysis of metalanguage (see Jaworski et al. 2004; Culpeper 2009; Busse and Hübler 2012), that is to say the search does not aim directly at the speech act in question but at passages where the speech act is an explicit topic in the ongoing discourse. Such a search uncovers cases in which the narrator or the fictional characters use a metapragmatic expression in order to assess whether such a speech act has been uttered or should be uttered and so on. But the search also retrieves cases in which the metapragmatic expression is used performatively to actually carry out such a speech act. In these instances, the metapragmatic expression functions as an illocutionary force indicating device and would also have shown up in the search for IFIDs reported in the previous section. All the retrieved hits, or a representative sample, then need to be inspected and categorised manually. It is necessary, in each case, to consult the wider context in order to see whether the named speech act is explicitly mentioned. Those examples can then be used for the analysis and categorisation. The information given in the speech act itself and in the vicinity of the speech act may not always be sufficient for an unambiguous categorisation, and passages are often difficult to interpret without the knowledge of the entire text in which they occur. Thus, the analysis is always a balancing act between trying to investigate a sufficient number of hits and spending sufficient time for each retrieved hit to make sense of a sufficiently large context.

In the current investigation, the search string *apolog\** was used in order to retrieve a range of ways in which apologising behaviour can be referred to. Table 3 gives the frequency of the most frequent forms that this string retrieved.

Apology	3,279
Apologize	4,089
Apologized	1,497
Apologies	1,206
Apologetically	1,099
Apologetic	2,091
Apologizing	598
Total	13,859

Table 3: Different forms of the search string *apolog\** and their overall frequency in COHA (fiction only)<sup>3</sup>

Again, only the fiction section of the *Corpus of Historical American English* (COHA) was used, and the corpus was split into four half-centuries. The retrieved hits were inspected manually. For each hit, it had to be decided whether it contained a relevant clearly identifiable apology, and if so, what kind of strategy or strategies were used. Interrater reliability checks were carried out on samples of 100 hits to ascertain a thorough understanding of the categories and an equal application of them to the data. In this case

<sup>3</sup> These figures again deviate slightly from those obtained through the official COHA website (see footnote 2 above).

several rounds were needed until a level of more than 70 per cent agreement between the coders was reached. Extensive discussions of problematic cases and mutual agreements on how to deal with them helped to improve the results and in the last round, a level of 77 per cent was reached. In a first step of the actual analysis the two coders proceeded to analyse a sufficient number of hits to reach 100 analysable apologies per half-century. Coder 1 coded periods A and C, and coder 2 coded periods B and D. The results obtained in this way – in spite of the careful interrater reliability check – revealed patterns that suggested some coder bias.<sup>4</sup> It was, therefore, decided to extend the number of hits to reach 200 analysable apologies for each period. This time coder 1 coded periods B and D, while coder 2 coded periods A and C. These are the results that I am going to present in the following.

On the basis of the entire number of hits for each period and the number of hits that were needed to identify 200 analysable ones, the approximate total number of analysable apologies for each period was calculated. For the first period, no calculations were needed because all retrieved hits were analysed and only 174 analysable ones were identified. The number of theoretically analysable apologies was then set in relation to the total number of words for this period. These figures are shown in Figure 3.

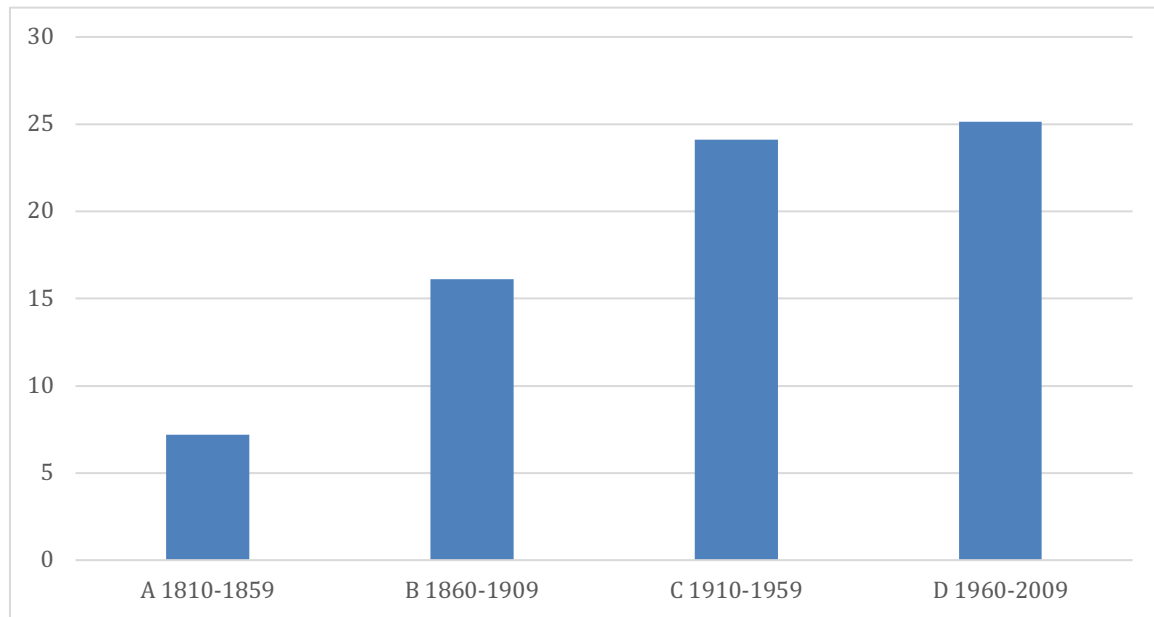


Figure 3: Approximate number of analysable apologies in COHA in a metapragmatic expression search with the search term *apolog\** (per million words)

The steady increase of the frequency of analysable apologies in Figure 3 accords well with a similar increase of the overall frequency of apology expressions shown in Figure 2 above. In Figure 3 the overall increase over the four periods is bigger (more than three-fold) while in Figure 2 the increase was about two-fold. Moreover, Figure 2 shows a significant

<sup>4</sup> Both coders found a relatively significant increase of analysable apologies in their two half centuries (for coder 1 from A to C, and for coder 2 from B to D) but coder 2 must have been somewhat more reluctant to recognize hits as analyzable apologies, which produced slight decreases from A to B and from C to D.

increase between period C and D, while Figure 3 shows a steady increase over the first three periods with only a slight further increase to period D. But it must be borne in mind that the two figures represent very different ways of tracing apologies. Neither of them gives a complete picture. Figure 2 focuses on those apologies that include an apology expression and should give a relatively complete account of these but it leaves out all those apologies that do not integrate an apology expression. Figure 3, on the other hand, includes only those apologies that somehow get explicitly named in the vicinity of their occurrence. Some of them include an IFID, and therefore overlap with those represented in Figure 2. Others do not. Together, however, the two figures provide clear evidence that the frequency with which apologies are mentioned in American English fictional texts has significantly increased over the last two centuries.

For the analysis of the analysable apologies a categorisation scheme was used that is based on the one proposed by Blum-Kulka et al. (1989a: 289-294) (see also Olshtain 1989). Table 4 provides an overview of the categories used for the analysis.

Category	Description
Illocutionary Force Indicating Device (IFID)	Elements such as <i>sorry</i> , <i>excuse</i> , <i>apologize</i> , etc.
Taking on Responsibility (RESP)	Speaker indicates explicitly that he/she carried out/was responsible for the offence
Denial of Intent (DINT)	Speaker claims that offence happened without his/her intention
Concern for hearer (CONH)	Speaker takes cognizance of the hearer's feelings
Explanation or Account (EXPL)	Speaker provides an explanation why the offence happened
Offer of Repair (REPR)	Speaker suggests a remedy for the offence
Promise of Forbearance (FORB)	Speaker promises that the offence will not happen again in the future
Non-verbal (NONV)	Apology is carried out without words

Table 4: Apology strategies (partly based on Blum-Kulka et al.'s 1989: 289-294)

Blum-Kulka et al.'s (1989a: 289-294) is much more fine-grained and offers many subcategories for these main categories. However, with our corpus-derived data such a detailed analysis turned out not to be feasible. Many of the subcategories were rare and would not have lent themselves for further statistical processing, or they proved too fuzzy to get through our interrater reliability check. Blum-Kulka et al. treat the category "Denial of Intent" as a subcategory of "Taking on Responsibility". I decided to treat it as an individual category because of its repeated presence in the data. I added a category of "Non-verbal" for non-verbal behaviour explicitly described as apologetic by or in the vicinity of the retrieved metapragmatic expression. I will illustrate all the categories below. Figure 4 gives an overview of the development of the different strategies. The figures are the percentages of the use of a specific strategy over all 200 retrieved hits for each half-century



(174 hits for the first half-century). In many cases, speakers use two or even more strategies in the same apology. These strategies are counted individually. For this reason, all the bars in Figure 4 add up to considerably more than 100 per cent.

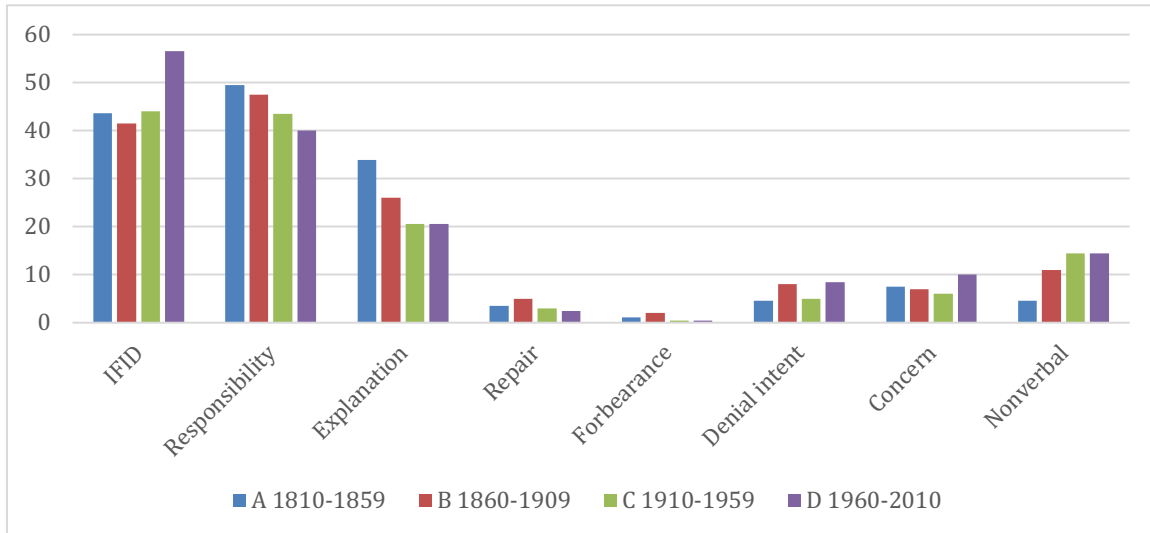


Figure 4: Percentage of strategy selection across four half-centuries (COHA Fiction)

Some distinct developments can be discerned from Figure 4. It is clear that the most important categories over all periods are IFIDs, Responsibility and Explanation. IFIDs account for more than 40 per cent of all analysed apologies in the first three periods and reach almost 60 per cent in the most recent one. Responsibility and Explanation decrease steadily and lose importance. In the most recent period they are only used in about 40 and 20 per cent of all the analysable apologies in the data. The other categories, in comparison, are relatively insignificant. There is a slight increase for Denial of Intent and a somewhat more marked increase for Non-verbal, which reaches about 15 per cent in the third and fourth period. Non-verbal is, of course, the only category that never co-occurs with any of the other categories. Whenever one of the other strategies is chosen, the apology cannot be nonverbal.

The category IFID was further analysed according to the type of IFID. This analysis overlaps somewhat with the analysis presented in section 5.2 above, but in this case the retrieval technique was very different and, therefore, it is not unexpected that the results differ somewhat. Figure 5 shows the development of the apology expressions that were attested in the set of analysed apologies.

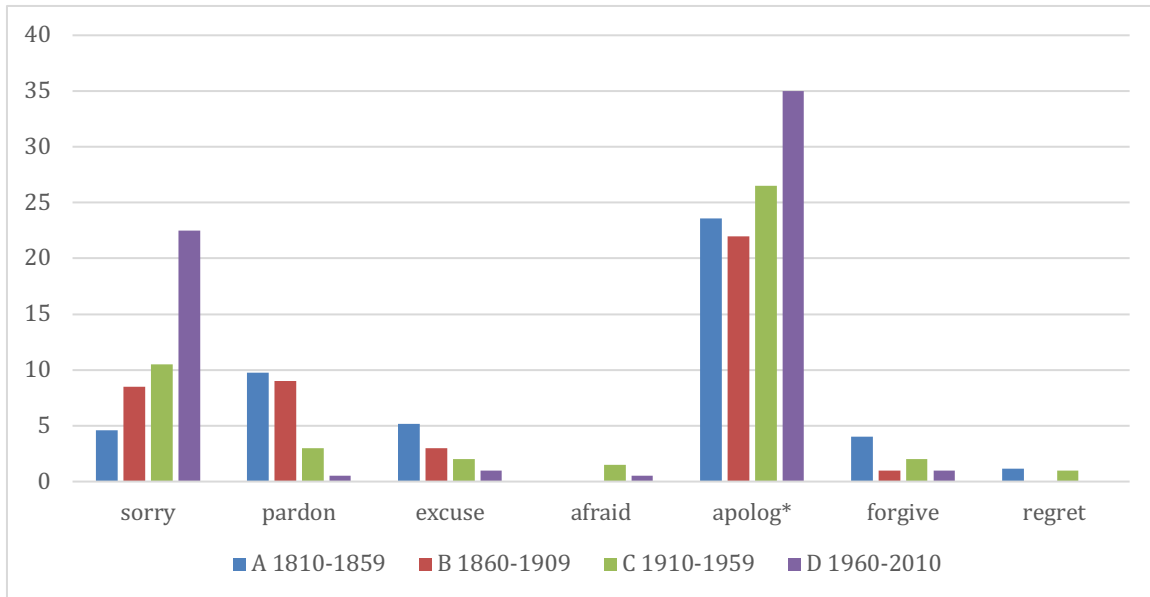


Figure 5: Percentage of apology expressions across four half-centuries (COHA Fiction)

Figure 5 shows interesting similarities and differences to Figure 1 above. Both figures indicate a very clear and steady increase of *sorry* and a clear decrease of *pardon*. *Forgive* is slightly more important in Figure 1 than in Figure 5 but in both cases, there appears to be a decrease over the periods. The apology expression *excuse*, on the other hand, shows a clear difference. In Figure 1 there is an increase over the four periods. In Figure 5, there is a decrease. The apology expression *apolog\** (subsuming the expressions *apology*, *apologize*, *apologetic* and so on) stands out. In Figure 1 it is almost non-existent, while in Figure 5 it shows frequencies between 20 and 35 per cent. The reason for this is the search technique used for the data in Figure 5. In Figure 1 all instances of *sorry* used as an IFID, for example, are represented (according to the limited precision of the used methods and calculations). In Figure 5, only those are included that occur near the metapragmatic expression *apolog\**. This is presumably a relatively small subset of all cases. In the case of *apolog\**, on the other hand, both figures should include all instances since the metapragmatic expression *apolog\** used as a search term automatically also retrieves all instances of the IFID *apolog\** (see examples below). The search term that retrieved the example is highlighted in bold, the identified strategies are highlighted in italics.

- (18) “I’m *sorry* – ” he began to **apologize**, but the old lady silenced him gracefully,  
(fic\_1852\_1850s\_9232)<sup>5</sup>
- (19) “I beg your *pardon*, sir,” said Paul, who, in spite of his desire to overtake Mike, felt it incumbent upon him to stop and offer an **apology**. “What do you mean, sir,” exploded the fat man, at last, “by tearing through the streets like a locomotive? You’ve nearly killed me.” “I am very *sorry*, sir.” “You ought to @ @ @ @ @ @

<sup>5</sup> References are taken from the downloaded Excel sheets with all retrieved hits from COHA. They indicate the genre (fiction in all cases), the year of publication, the relevant decade and an identification number). They can most easily be located in the corpus by using an identifying string from the extract as a search term in COHA.

@ @ @ @ at such speed? You ought to be indicted as a public nuisance.” “*I was trying to catch a thief*,” said Paul. (fic\_1871\_1870s\_247)<sup>6</sup>

- (20) Mr Odendaal, I **apologize** for turning up without telephoning.  
(fic\_1993\_1990s\_25005)
- (21) “Mrs. Carroll has gone into the dining-room,” the servant told them at the door, and Sydney assumed much cheerfulness as she made her **apologies**. “I’ve brought Bob, grandmother. He’s been all over everywhere with me this morning. You’ll *forgive* me, Katrina, for leaving you, won’t you? Where’s Mr. Wendell?”  
(fic\_1903\_1900s\_4763)

Extracts (18) to (21) illustrate a range of different apologies that contain an IFID, and they illustrate some of the problems that can be encountered in a manual analysis of retrieved hits. In (18) the IFID *sorry* is used. Extract (19) shows that a careful look is required from the coders. In this case the recorded apology extends over several lines. The speaker first apologises with “I beg your pardon” and then adds “I’m very sorry” and – after an intervention by his interlocutor – adds an explanation, “I was trying to catch a thief”, which is yet another apology strategy. In extract (20) the metapragmatic expression *apolog\** happens to be an IFID itself, in the form of *apologize*. Extract (21), finally, illustrates one of the few cases in which *forgive* is used as an IFID.

Extract (22) to (25) illustrate the apology strategies “Taking responsibility” and “Explanation”.

- (22) *I offer this as an **apology** for not prefixing to this book, according to custom, half a dozen pages of useless matter, like a clumsy, ostentatious vestibule to a house that would be more easily entered without one.* (fic\_1823\_1820s\_7213)
- (23) “*I was woolgathering*,” he said by way of **apology**. (fic\_1971\_1970s\_10537)
- (24) “It is certainly not a personal letter,” said Bess, maliciously glancing at the superscription. “Don’t you see it is addressed to ‘Mrs. Glenn and daughters.’” “*In a time like that people don’t think much of letters*,” commented Mrs. Glenn, **apologetically**. (fic\_1889\_1880s\_5842)
- (25) Eddie drank a gulp or two of punch, craning his neck in a pretense of looking for his friends. “*It’s awful hard to find anybody in a crowd like this*,” he **apologized**. “Either that or the band boys have beat it.” (fic\_1930\_1930s\_23157)

In extracts (22) and (23), the respective speakers offer what is explicitly described as an apology. Without this designation it might have been difficult for a coder or researcher to identify an apology at all, at least without considering much more context. In both cases what the speaker says might also have been categorised as an explanation, but as a rule, explanations that include the personal pronoun *I* as a responsible part of what happened were categorized as the strategy “Taking on responsibility”. In the explanations given in

---

<sup>6</sup> Series of @-signs are interspersed in the corpus for copyright reasons. Apparently, they were considered to be unproblematic for corpus searches by the corpus compilers, but for the pragmaticist who tries to manually code retrieved hits they can be a serious hindrance in understanding the text.

(24) and (25), on the other hand, reasons external to the speaker are given as an explanation for why the offence happened. Such cases were categorised as Explanations.

Extracts (26) to (30), finally, illustrate the remaining five strategies that are attested considerably less frequently.

- (26) She paused a moment, and then, in an **apologetic** tone, she added, “*I’d be perfectly willing to talk with you about it generally*, my dear Dorothy, but not now.” (fic\_1896\_1890s\_948)
- (27) “We do not use compliments, Richard,” said he; “my daughter’s name is Asenath.” I beg *pardon*. *I will try to accustom myself to your ways*, since you have been so kind as to take me for a while,” **apologized** Richard Hilton. (fic\_1872\_1870s\_95)
- (28) “We haven’t even discussed what we’re going to do this evening,” she said, adding **apologetically**, “*I didn’t mean to monopolize the conversation.*” (fic\_1986\_1980s\_10741)
- (29) “My dearest child,” it read. “My instructions are that you shall read this only if I die. Grumbach is a good fellow and Canon Masson is my friend. Trust them. *I am sorry I have not given you more reason to trust me*, but I shall not bore you with regrets or **apologies**. (fic\_1977\_1970s\_780162)
- (30) Morley *nodded* **apologetically**, but before he could answer Sorenson pushed him away. (fic\_1934\_1930s\_10048)

In extract (26) the speaker offers what appears to be a repair for an earlier offence. The fact that this is meant as an apology is signalled by the narrator’s description of her tone of voice as apologetic. The apology in extract (27) uses two strategies. It combines the IFID *pardon* with a promise of forbearance. The speaker promises to refrain from complimenting in future as this appears not to be the custom of his interlocutors. Extract (28) illustrates a speaker who denies any intention of committing what appears to have been the offence. Here, too, the utterance is signalled as an apology through the narratorial voice. Extract (29) is one of the relatively few examples of the strategy Concern for the Hearer. It is here combined with the IFID *sorry*. Extract (30), finally, is an interesting case. There is no real apology, at least not a verbal one. One of the characters is said to have nodded apologetically. Thus, the gesture itself is described by the narrator as an apology.

Extract (31) is a particularly rich example. It includes the use of several different apology strategies.

- (31) Rightfully so, she told herself, and she **apologized**, though it was a double humiliation to do so. “*I was terribly wrong to sneer at your lack of knowledge of science and at your mistaken beliefs*,” she said. “*It is not your fault that you were born in 1619, and I should not have taunted you with that*. I did so just to make you so mad I’d get an edge on you. It was a rotten thing to do. *I promise not to do it again*, and I most abjectly beg your *pardon*. *I did not really mean it.*” (fic\_1977\_1970s\_10627)

The speaker starts off with taking responsibility for the offence (“I was terribly wrong ...”); she gives an explanation (“it is not your fault that ...”), takes responsibility once more (“I should not have taunted you”), adds a promise of forbearance (“I promise not to do it

again”) followed by an intensified IFID (“I most abjectly beg your pardon”), and at the very end finishes off with a denial of intent (“I did not really mean it”). The overall effect seems to be one of comic exaggeration and probably bears more than just a hint of irony or even sarcasm, but without more context the precise nature of this is hard to ascertain.

#### 4 Discussion and conclusion

It has to be stressed again that the data analysed here does not consist of everyday spontaneous conversations. It consists exclusively of fictional data and fictional representations of conversations because in this type of data, apologies turn out to be much more frequent than in other types of data that are available in historical corpora, such as magazines, newspapers or non-fiction books. As pointed out above, fictional texts are here not presented as a substitute of everyday spoken interaction, they are something clearly different (see Jucker and Locher 2017 for a detailed discussion).

The analysis presented above has combined two different methods of retrieving apologies from the *Corpus of Historical American English* in order to gain a more comprehensive perspective on their diachronic development. On the one hand, a search was carried out for illocutionary force indicating devices (IFIDs) and on the other hand, the metapragmatic expression *apology* and its derivatives were used to retrieve passages in which apologies were mentioned explicitly. In the former case, it was necessary to manually remove false hits, and because of the great number of hits this had to be done on the basis of representative samples of hits with subsequent extrapolation of the figures to the entire corpus. In the latter case, the coders first had to manually extract the analysable apologies from all the hits retrieved with the metapragmatic expression *apolog\** in order to subject these to a detailed analysis of the strategies employed to carry out the apology.

A fine-grained analysis of all the retrieved apologies turned out to be very difficult. It was not possible, for instance, to consistently code different offence categories. Even with a severely reduced set of categories inter-rater reliability tests failed. On the other hand, it was possible to distinguish between different strategies of performing an apology, such as IFID, Taking on Responsibility, Explanation, and so on, but it was not possible to use such subtle sub-categories as had been employed in experimental research (e.g. Blum-Kulka et al. 1989a, Trosborg 1995) or the even more detailed categories used in Deutschmann’s (2003) corpus-based analysis. It is, of course, possible that apologies elicited in DCTs or in role plays and apologies in conversational data are easier to analyse than representations of apologies in fictional texts with the various levels of narratorial voice and depicted characters.

In spite of the limited nature of the data, both analyses provided interesting and specific results which reinforce each other. The frequency of apologies containing an illocutionary force indicating device steadily increased from about one hundred per million words in the first half century to over two hundred per million words in the last half century of the period under investigation. The metapragmatic expression analysis produced a similar increase of analysable apologies from 7 to 25 instances per million words. Together the two investigations provide very strong empirical evidence for a significant increase in the overall frequency of apologies in the fiction section of the *Corpus of Historical American English* even though both investigations on their own can only offer an incomplete picture.

It is interesting to speculate why apologies should have become so much more important over the last two hundred years. It was clearly not a sudden increase at a certain point in the history but a gradual increase. Perhaps the increase in frequency goes together with a decrease in the weight of an apology. What used to be a heartfelt expression of regret for having committed an offence has in many cases turned to a conventionalized phrase with little meaning.

The massive increase of *sorry*, which has become by far the most important apology expression today, supports such an interpretation. At the beginning of the nineteenth century, it was one among several such expressions, notably *pardon* and *forgive*. In the second half of the twentieth century, it outnumbers the total of all the other expressions by about three to one. *Pardon* and *forgive* can both be understood as appeals to the addressee to exonerate the speaker from some more or less serious wrongdoing, while *sorry* is more an expression of regret (see also Jucker and Taavitsainen 2008, where we offer similar speculations on the basis of data from the sixteenth and seventeenth centuries). *Sorry* avoids the imposition on the addressee to absolve the speaker from blame and merely serves as a token acknowledgment of some minor problem caused by the speaker.

The development of the selection of apology strategies is less clear-cut but in some ways it also supports this interpretation. The two categories Taking on Responsibility and Offering an Explanation both steadily decreased over the four half centuries while IFIDs and non-verbal apologies increased in their frequency. Apologies appear to become gradually more routinized. Explanations are no longer needed. A brief apology expression or even an apologetic nod or glance is enough to perform an apology, at least in the fictional data investigated in this project. It remains a matter of speculation as to whether these developments are restricted to fictional texts or whether they are signs of a wider – perhaps socio-cultural – development.

## Acknowledgment

First and foremost, my thanks must go to Nina Helg-Kurmann and Lukas Zbinden, Master students at the University of Zurich, who worked as my research assistants and carried out the majority of the coding work reported in this paper. They both spent many hours pondering extracts of texts from the COHA, trying to make sense of the interactions between the characters. I also offer heartfelt thanks to Magdalena Leitner and Mirjam Schmalz, who read a draft version of this paper and provided a wealth of insightful comments and suggestions for its improvement. The usual disclaimers apply.

## Data sources

Corpus of Contemporary American English. Official website: <https://corpus.byu.edu/coca/>

Corpus of Historical American English. Official website: <https://corpus.byu.edu/coha/>

Corpus of Historical American English. Accessed through the Dependency Databank at the English Department, University of Zurich. <http://es-dbank.uzh.ch>

## References

Aijmer, Karin. (1996) Conversational Routines in English. Convention and Creativity. London: Longman.

- Blum-Kulka, Shoshana, Juliane House and Gabriele Kasper (eds.). (1989a) *Cross-Cultural Pragmatics: Requests and Apologies*. Norwood, NJ: Ablex.
- Blum-Kulka, Shoshana, Juliane House and Gabriele Kasper. (1989b) Investigating cross-cultural pragmatics: An introductory overview. In: Shoshana Blum-Kulka, Juliane House and Gabriele Kasper (eds.). *Cross-Cultural Pragmatics: Requests and Apologies*. Norwood, NJ: Ablex, 1-34.
- Brown, Penelope, and Stephen C. Levinson. (1987) *Politeness. Some Universals in Language Usage*. (Studies in Interactional Sociolinguistics 4). Cambridge: Cambridge University Press.
- Busse, Ulrich, and Axel Hübner (eds.). (2012) *Investigations into the Meta-Communicative Lexicon of English. A Contribution to Historical Pragmatics*. (Pragmatics & Beyond New Series 220). Amsterdam/Philadelphia: John Benjamins.
- Culpeper, Jonathan. (2009) The metalanguage of impoliteness: Using sketch engine to explore the Oxford English Corpus. In: Paul Baker (ed.). *Contemporary Corpus Linguistics*. (Contemporary Studies in Linguistics). London: Continuum, 64-86.
- Deutschmann, Mats (2003) *Apologising in British English*. (Skrifter från moderna språk 10). Umeå: Institutionen för moderna språk, Umeå University.
- Drew, Paul, Alexa Hepburn, Piera Margutti and Renata Galatolo. (2016) Introduction to the special issue on apologies in discourse. *Discourse Processes* 53.1-2, 1-4.
- Goffman, Erving. (1971) *Relations in Public: Microstudies of the Public Order*. New York: Harper and Row.
- Heritage, John, and Chase Wesley Raymond. (2016) Are explicit apologies proportional to the offenses they address? *Discourse Processes* 53.1-2, 5-25.
- Holmes, Janet. (1990) Apologies in New Zealand English. *Language in Society* 19.2, 155-199.
- Jaworski, Adam, Nikolas Coupland and Dariusz Galasiński (eds.). (2004) *Metalanguage. Social and Ideological Perspectives*. Berlin: Mouton de Gruyter.
- Jucker, Andreas H. (2017) Speech acts and speech act sequences: Greetings and farewells in the history of American English. *Studia Neophilologica*, 1-20.
- Jucker, Andreas H., and Miriam A. Locher (2017) Introducing Pragmatics of Fiction: Approaches, trends and developments. In: Miriam A. Locher and Andreas H. Jucker (eds.). *Pragmatics of Fiction* (Handbooks of Pragmatics 12). Berlin: de Gruyter, 1-21.
- Jucker, Andreas H., and Irma Taavitsainen (2008) Apologies in the history of English: Routinized and lexicalized expressions of responsibility and regret. In: Andreas H. Jucker and Irma Taavitsainen (eds.). *Speech Acts in the History of English*. (Pragmatics & Beyond New Series 176). Amsterdam/Philadelphia: John Benjamins, 229-244.
- Jucker, Andreas H., and Irma Taavitsainen. (2013) *English Historical Pragmatics*. (Edinburgh Textbooks on the English Language). Edinburgh: Edinburgh University Press.
- Jucker, Andreas H., and Irma Taavitsainen. (2014) Complimenting in the history of American English: A metacommunicative expression analysis. In: Irma Taavitsainen, Andreas H. Jucker and Jukka Tuominen (eds.). *Diachronic Corpus Pragmatics*. (Pragmatics & Beyond New Series 243). Amsterdam: John Benjamins, 257-276.
- Jucker, Andreas H., Gerold Schneider, Irma Taavitsainen and Barb Breustedt. (2008) Fishing for compliments: Precision and recall in corpus-linguistic compliment research. In: Andreas H. Jucker, and Irma Taavitsainen (eds.). *Speech Acts in the History of English*. (Pragmatics & Beyond New Series 176). Amsterdam/Philadelphia: John Benjamins, 273-294.
- Kohnen, Thomas. (2017) Non-canonical speech acts in the history of English. *Zeitschrift für Anglistik und Amerikanistik* 65.3, 303-318.

- Lakoff, Robin Tolmach. (2001) Nine ways of looking at apologies: The necessity for interdisciplinary theory and method in discourse analysis. In: Deborah Schiffrin, Deborah Tannen and Heidi E. Hamilton (eds.). *The Handbook of Discourse Analysis*. Oxford: Blackwell, 199-214.
- Lutzky, Ursula, and Andrew Kehoe. (2017a) "Oops, I didn't mean to be so flippant": A corpus pragmatic analysis of apologies in blog data. *Journal of Pragmatics* 116, 27-36.
- Lutzky, Ursula, and Andrew Kehoe. (2017b) "I apologise for my poor blogging": Searching for apologies in the *Birmingham Blog Corpus*. *Corpus Pragmatics* 1, 37-56.
- Manes, Joan, and Nessa Wolfson. (1981) The compliment formula. In: Florian Coulmas (ed.). *Conversational Routine. Explorations in Standardized Communication Situations and Prepatterned Speech*. The Hague: Mouton, 115-132.
- Meier, A. J. (1998) Apologies: What do we know? *International Journal of Applied Linguistics* 8.2, 215-231.
- Ogierman, Eva. (2018) Discourse completion tasks. In: Andreas H. Jucker, Wolfram Bublitz and Klaus P. Schneider (eds.). *Methods in Pragmatics*. (Handbooks of Pragmatics 10). Berlin: de Gruyter.
- Olshtain, Elite. (1989) Apologies across languages. In: Shoshana Blum-Kulka, Juliane House, and Gabriele Kasper (eds.). *Cross-Cultural Pragmatics: Requests and Apologies*. Norwood, NJ: Ablex, 155-173.
- Owen, Marion. (1983) *Apologies and Remedial Interchanges. A Study of Language Use in Social Interaction*. Berlin: Mouton.
- Robinson, Jeffrey D. (2004) The sequential organization of "explicit" apologies in naturally occurring English. *Research on Language and Social Interaction* 37.3, 291-330.
- Searle, John R. (1969) *Speech Acts. An Essay in the Philosophy of Language*. Cambridge: Cambridge University Press.
- Trosborg, Anna. (1987) Apology strategies in natives / non-natives. *Journal of Pragmatics* 11, 147-167.
- Trosborg, Anna. (1995) *Interlanguage Pragmatics. Requests, Complaints and Apologies*. Berlin: Mouton de Gruyter.